IN THE CLAIMS:

Status of the claims:

Claims 24 is currently amended;

Claims 1-23 and 25-75 are original.

Please amend the claims as follows:

1.   (original) A system for generating a description record from video information, comprising:

at least one video input interface for receiving said video information;

a computer processor coupled to said at least one video input interface for receiving said video information therefrom, processing said video information by performing video object extraction processing to generate video object descriptions from said video information, processing said generated video object descriptions by object hierarchy construction and extraction processing to generate video object hierarchy descriptions, and processing said generated video object descriptions by entity relation graph generation processing to generate entity relation graph descriptions, wherein at least one description record including said video object descriptions, said video object hierarchy descriptions and said entity relation graph descriptions is generated to represent content embedded within said video information; and

a data storage system, operatively coupled to said processor, for storing said at least one description record.

2.   (original)   The system of claim 1, wherein said video object extraction

processing and said object hierarchy construction and extraction processing are performed in parallel.

3. (original) The system of claim 1, wherein said video object extraction processing comprises: video segmentation processing to segment each video in said video information into regions within said video; and

feature extraction and annotation processing to generate one or more feature descriptions for one or more said regions;

whereby said generated video object descriptions comprise said one or more feature descriptions for one or more said regions.

4. (original) The system of claim 3, wherein said regions are selected from the group consisting of local, segment and global regions.

5. (original) The system of claim 3, wherein said one or more feature descriptions are selected from the group consisting of media features, visual features, temporal features, and semantic features.

6. (original) The system of claim 5, wherein said semantic features are further defined by at least one feature description selected from the group consisting of who, what object, what action, where, when, why, and text annotation.

7. (original) The system of claim 5, wherein said visual features are further defined

by at least one feature description selected from the group consisting of color, texture, position, size, shape, motion, camera motion, editing effect, and orientation.

8. (original) The system of claim 5, wherein said media features are further defined by at least one feature description selected from the group consisting of file format, file size, color representation, resolution, data file location, author, creation, scalable layer and modality transcoding.

9. (original) The system of claim 5, wherein said temporal features are further defined by at least one feature description selected from the group consisting of start time, end time and duration.

10. (original) The system of claim 1, wherein said object hierarchy construction and extraction processing generates video object hierarchy descriptions of said video object descriptions based on visual feature relationships of video objects represented by said video object descriptions.

11. (original) The system of claim 1, wherein said object hierarchy construction and extraction processing generates video object hierarchy descriptions of said video object descriptions based on semantic feature relationships of video objects represented by said video object descriptions.

12. (original) The system of claim 1, wherein said object hierarchy construction and

extraction processing generates video object hierarchy descriptions of said video object descriptions based on media feature relationships of video objects represented by said video object descriptions.

13. (original) The system of claim 1 , wherein said object hierarchy construction and extraction processing generates video object hierarchy descriptions of said video object descriptions based on relationships of video objects represented by said video object descriptions, wherein said relationships are selected from the group consisting of visual feature relationships, semantic feature relationships, temporal feature relationships and media feature relationships.

14. (original) The system of claim 1, wherein said object hierarchy construction and extraction processing generates video object hierarchy descriptions of said video object descriptions based on relationships of video objects represented by said video object descriptions, wherein said video object hierarchy descriptions have a plurality of hierarchical levels.

15. (original) The system of claim 14, wherein said video object hierarchy descriptions having a plurality of hierarchical levels comprise clustering hierarchies.

16. (original) The system of claim 15, wherein said clustering hierarchies are based on relationships of video objects represented by said video object descriptions, wherein said relationships are selected from a group consisting of visual feature relationships,

semantic feature relationships, temporal relationships and media feature relationships.

17. (original) The system of claim 15, wherein said video object hierarchy descriptions having a plurality of hierarchical levels are configured to comprise multiple levels of abstraction hierarchies.

18. (original) The method of claim 17, wherein said multiple levels of abstraction hierarchies are configured to be based on relationships of video objects represented by said video object descriptions, wherein said relationships are selected from a group consisting of visual feature relationships, semantic feature relationships, temporal feature relationships and media feature relationships.

19. (original) The system of claim 1, wherein said entity relation graph generation processing generates entity relation graph descriptions of said video object descriptions based on relationships of video objects represented by said video object descriptions, wherein said relationships are selected from the group consisting of visual feature relationships, semantic feature relationships, temporal feature relationships and media feature relationships.

20. (original) The system of claim 1, further comprising an encoder for receiving and encoding said video object descriptions into encoded description information, wherein said data storage system is operative to store said encoded description information as said at least one description record.

21. (original) The system of claim 1, wherein said video object descriptions, said video object hierarchy descriptions, and said entity relation graph descriptions are combined together to form video descriptions, and further comprising an encoder for receiving and encoding said video descriptions into encoded description information, wherein said data storage system is operative to store said encoded description information as said at least one description record.

22. (original) The system of claim 21, wherein said encoder comprises a binary encoder.

23. (original) The system of claim 21, wherein said encoder comprises an XML encoder.

24. (currently amended) The system of claim 1, further comprising:

a video display device operatively coupled to the computer processor for displaying the video information; and

at least one user input device operatively coupled to the computer processor, wherein at least a portion of said ~~video object processing~~ video object extraction processing, said object hierarchy construction and extraction processing, or said entity relation graph generation processing includes receiving a user input through manipulation of said user input device.

25. (original) A method for generating a description record from video information, comprising the steps of:

receiving said video information;

processing said video information by performing video object extraction processing to generate video object descriptions from said video information;

processing said generated video object descriptions by object hierarchy construction and extraction processing to generate video object hierarchy descriptions;

processing said generated video object descriptions by entity relation graph generation processing to generate entity relation graph descriptions, wherein at least one description record including said video object descriptions, said video object hierarchy descriptions and said entity relation graph descriptions is generated to represent content embedded within said video information; and

storing said at least one description record.

26. (original) The method of claim 25, wherein said steps of video object extraction processing and object hierarchy construction and extraction processing are performed in parallel.

27. (original) The method of claim 25, wherein said step of video object extraction processing comprises the further steps of:

video segmentation processing to segment each video in said video information into regions within said video; and

feature extraction and annotation processing to generate one or more feature

descriptions for one or more said regions;

whereby said generated video object descriptions comprise said one or more feature descriptions for one or more said regions.

28. (original) The method of claim 27, wherein said regions are selected from the group consisting of local, segment and global regions.

29. (original) The method of claim 27, further comprising the step of selecting said one or more feature descriptions from the group consisting of media features, visual features, temporal and semantic features.

30. (original) The method of claim 29, wherein said semantic features are further defined by at least one feature description selected from the group consisting of who, what object, what action, where, when, why and text annotation.

31. (original) The method of claim 29, wherein said visual features are further defined by at least one feature description selected from the group consisting of color, texture, position, size. shape, motion, editing effect, camera motion and orientation.

32. (original) The method of claim 29, wherein said media features are further defined by at least one feature description selected from the group consisting of file format, file size, color representation, resolution, data file location, author, creation, scalable layer and modality transcoding.

33. (original) The method of claim 29, wherein said temporal features are further defined by at least one feature description selected from the group consisting of start time, end time and duration.

34. (original) The method of claim 25, wherein said step of object hierarchy construction and extraction processing generates video object hierarchy descriptions of said video object descriptions based on visual feature relationships of video objects represented by said video object descriptions.

35. (original) The method of claim 25, wherein said step of object hierarchy construction and extraction processing generates video object hierarchy descriptions of said video object descriptions based on semantic feature relationships of video objects represented by said video object descriptions.

36. (original) The method of claim 25, wherein said step of object hierarchy construction and extraction processing generates video object hierarchy descriptions of said video object descriptions based on media feature relationships of video objects represented by said video object descriptions.

37. (original) The method of claim 25, wherein said step of object hierarchy construction and extraction processing generates video object hierarchy descriptions of said video object descriptions based on temporal feature relationships of video objects

represented by said video object descriptions.

38. (original) The method of claim 25, wherein said step of object hierarchy construction and extraction processing generates video object hierarchy descriptions of said video object descriptions based on relationships of video objects represented by said video object descriptions, wherein said relationships are selected from the group consisting of visual feature relationships, semantic feature relationships, temporal feature relationships and media feature relationships.

39. (original) The method of claim 25, wherein said step of object hierarchy construction and extraction processing generates video object hierarchy descriptions of said video object descriptions based on relationships of video objects represented by said video object descriptions, wherein said video object hierarchy descriptions are configured to have a plurality of hierarchical levels.

40. (original) The method of claim 39, wherein said video object hierarchy descriptions having a plurality of hierarchical levels are configured to comprise clustering hierarchies.

41. (original) The method of claim 40, wherein said clustering hierarchies are configured to be based on relationships of video objects represented by said video object descriptions, wherein said relationships are selected from a group consisting of visual feature relationships, semantic feature relationships, temporal feature relationships and

media feature relationships.

42. (original)   The method of claim 40, wherein said video object hierarchy descriptions having a plurality of hierarchical levels are configured to comprise multiple levels of abstraction hierarchies.

43. (original) The method of claim 40, wherein said multiple levels of abstraction hierarchies are configured to be based on relationships of video objects represented by said video object descriptions, wherein said relationships are selected from a group consisting of visual feature relationships, semantic feature relationships, temporal feature relationships and media feature relationships.

44. (original) The method of claim 25, wherein said step of entity relation graph generation processing generates entity relation graph descriptions of said video object descriptions based on relationships of video objects represented by said video object descriptions, wherein said relationships are selected from the group consisting of visual feature relationships, semantic feature relationships, temporal feature relationships and media feature relationships.

45. (original) The method of claim 25, further comprising the steps of receiving and encoding said video object descriptions into encoded description information, and storing said encoded description information as said at least one description record.

46. (original) The method of claim 25, further comprising the steps of combining said video object descriptions, said video object hierarchy descriptions, and said entity relation graph descriptions to form video descriptions, and receiving and encoding said video descriptions into encoded description information, and storing said encoded description information as said at least one description record.

47. (original) The method of claim 46, wherein said step of encoding comprises the step of binary encoding.

48. (original) The method of claim 46, wherein said step of encoding comprises the step of XML encoding.

49. (original) A computer readable media containing digital information with at least one description record representing video content embedded within corresponding video information, the at least one description record comprising:

one or more video object descriptions generated from said video information using video object extraction processing;

one or more video object hierarchy descriptions generated from said generated video object descriptions using object hierarchy construction and extraction processing; and

one or more entity relation graph descriptions generated from said generated video object descriptions using entity relation graph generation processing.

50. (original) The computer readable media of claim 49, wherein said video object

descriptions, said video object hierarchy descriptions, and said entity relation graph descriptions further comprise one or more feature descriptions.

51. (original) The computer readable media of claim 50, wherein said one or more feature descriptions are selected from the group consisting of media features, visual features, temporal features and semantic features.

52. (original) The computer readable media of claim 51, wherein said semantic features are further defined by at least one feature description selected from the group consisting of who, what object, what action, where, when, why and text annotation.

53. (original) The computer readable media of claim 51, wherein said visual features are further defined by at least one feature description selected from the group consisting of color, texture, position, size, shape, motion, camera motion, editing effect and orientation.

54. (original) The computer readable media of claim 51, wherein said media features are further defined by at least one feature description selected from the group consisting of file format, file size, color representation, resolution, data file location, author, creation, scalable layer and modality transcoding.

55. (original) The computer readable media of claim 51, wherein said temporal features are further defined by at least one feature description selected from the group

consisting of start time, end time and duration.

56. (original) The computer readable media of claim 49, wherein said object hierarchy descriptions are based on visual feature relationships of video objects represented by said video object descriptions.

57. (original) The computer readable media of claim 49, wherein said video object hierarchy descriptions are based on semantic feature relationships of video objects represented by said video object descriptions.

58. (original) The computer readable media of claim 49, wherein said video object hierarchy descriptions are based on media feature relationships of video objects represented by said video object descriptions.

59. (original) The computer readable media of claim 49, wherein said video object hierarchy descriptions are based on temporal feature relationships of video objects represented by said video object descriptions.

60. (original) The computer readable media of claim 49, wherein said video object hierarchy descriptions are based on relationships of video objects represented by said video object descriptions, wherein said video object hierarchy descriptions have a plurality of hierarchal levels.

61. (original) The computer readable media of claim 60, wherein said video object hierarchy descriptions having a plurality of hierarchal levels comprise clustering hierarchies.

62. (original) The computer readable media of claim 61, wherein said clustering hierarchies are based on relationships of video objects represented by said video object descriptions, wherein said relationships are selected from a group consisting of visual feature relationships, semantic feature relationships, temporal feature relationships and media feature relationships.

63. (original) The computer readable media of claim 62, wherein said video object hierarchy descriptions having a plurality of hierarchical levels are configured to comprise multiple levels of abstraction hierarchies.

64. (original) The computer readable media of claim 63, wherein said multiple levels of abstraction hierarchies are configured to be based on relationships of video objects represented by said video object descriptions, wherein said relationships are selected from a group consisting of visual feature relationships, semantic feature relationships, temporal feature relationships and media feature relationships.

65. (original) The computer readable media of claim 49, wherein said entity relation graph descriptions are based on relationships of video objects represented by said video object descriptions, wherein said relationships are selected from the group consisting of

visual feature relationships, semantic feature relationships, temporal feature relationships and media feature relationships.

66. (original) The computer readable media of claim 49, wherein said video object descriptions are in the form of encoded description information.

67. (original) The computer readable media of claim 49, wherein said video object descriptions, said video object hierarchy descriptions, and said entity relation graph descriptions are combined together in the form of encoded description information.

68. (original) The computer readable media of claim 67, wherein said encoded description information is in the form of binary encoded information.

69. (original) The computer readable media of claim 67, wherein said encoded description information is in the form of XML encoded information.

70. (original) The system of claim 1, wherein feature descriptions include pointers to extraction and matching code to facilitate code downloading.

71. (original) The system of claim 5, wherein feature descriptions include pointers to extraction and matching code to facilitate code downloading.

72. (original) The method of claim 25, wherein feature descriptions include pointers

to extraction and matching code to facilitate code downloading.

73. (original) The method of claim 29, wherein feature descriptions include pointers to extraction and matching code to facilitate code downloading.

74. (original) The computer readable media of claim 49, wherein feature descriptions include pointers to extraction and matching code to facilitate code downloading.

75. (original) The computer readable media of claim 53, wherein feature descriptions include pointers to extraction and matching code to facilitate code downloading.